

03-10-06

IN THE U.S. PATENT AND TRADEMARK OFFICE
Patent Application Transmittal Letter

ASSISTANT COMMISSIONER FOR PATENTS
Washington, D.C. 20231

JCS53 U.S. 09/10/2009
03/09/00

Sir:

Transmitted herewith for filing under 37 CFR 1.53(b) is a(n): Utility Design
 original patent application,
 continuation-in-part application

INVENTOR(S): Tony M. Brewer, et al

TITLE: Protocol For Insuring Exactly Once Semantics Of Transactions Across An Unordered, Unreliable Network

Enclosed are:

The Declaration and Power of Attorney. signed unsigned or partially signed
 2 sheets of drawings (one set) Associate Power of Attorney
 Form PTO-1449 Information Disclosure Statement and Form PTO-1449
 Priority document(s) (Other) (fee \$ _____)

CLAIMS AS FILED BY OTHER THAN A SMALL ENTITY				
(1) FOR	(2) NUMBER FILED	(3) NUMBER EXTRA	(4) RATE	(5) TOTALS
TOTAL CLAIMS	20 — 20	0	X \$18	\$ 0
INDEPENDENT CLAIMS	3 — 3	0	X \$78	\$ 0
ANY MULTIPLE DEPENDENT CLAIMS	0		\$260	\$ 0
BASIC FEE: Design \$310.00 ; Utility \$690.00)				\$ 690
TOTAL FILING FEE				\$ 690
OTHER FEES				\$
TOTAL CHARGES TO DEPOSIT ACCOUNT				\$ 690

Charge \$ 690 to Deposit Account 08-2025. At any time during the pendency of this application, please charge any fees required or credit any over payment to Deposit Account 08-2025 pursuant to 37 CFR 1.25. Additionally please charge any fees to Deposit Account 08-2025 under 37 CFR 1.16, 1.17, 1.19, 1.20 and 1.21. A duplicate copy of this sheet is enclosed.

"Express Mail" label no. EL483249633US

Date of Deposit 3/9/00

Respectfully submitted,

Tony M. Brewer, et al

By Leslie P Gehman

Leslie P Gehman

Attorney/Agent for Applicant(s)
Reg. No. P45,624

Date: March 9, 2000

Telephone No.: (970) 898-3642

**PROTOCOL FOR INSURING EXACTLY ONCE SEMANTICS OF TRANSACTIONS
ACROSS AN UNORDERED, UNRELIABLE NETWORK**

Tony M. Brewer
5227 Mariners Drive
Plano, Texas 75093
Citizenship: U.S.A.

Gregory S. Palmer
3012 Mason Drive
Plano, Texas 75025
Citizenship: U.S.A.

TECHNICAL FIELD

The invention relates in general to data transmission over a network and in particular to efficiently monitoring such data transmission over an unreliable network.

BACKGROUND

Generally when conducting communication including transmitting messages, packets, or transactions between data processing or storage devices, some form of protocol or handshaking is required to manage and coordinate the communication. Where only two devices are transmitting information back and forth, such communication management may be relatively straightforward. Generally, direct and continuous contact may be established between two devices enabling misdirected transmissions, omitted information, and erroneous data to be readily identified and corrected.

Such communication management generally becomes more difficult and complex where a plurality of sources communicate with a plurality of destinations across a network as is shown in FIGURE 1. In such a network, direct and continuous contact between a source and destination for a particular message will generally not be available. Instead the communication process will generally require that messages be transmitted via one or more intermediate points in the network in between transmission by the source and reception by the ultimate destination. A number of problems may be introduced at the intermediate points in the network including the loss of information, introduction of erroneous information, or a hardware failure causing a complete transmission failure. Such events may leave both the source and destination of the transmission unaware of the problem and unable to resolve it.

One solution employed in the prior art to deal with this problem is the association of sequence numbers with messages or packets transmitted from a source to a destination. With sequence numbers attached to each message, destination devices can check the sequence number associated with a message to ensure that it is the next message in sequence from the source location concerned. If a duplicate message is received at a destination device, the message is generally ignored. If a sequence of messages is received wherein a sequence number has been skipped, a protocol active between the destination and the source will generally cause the skipped message to be resent.

The use of sequence numbers for communication management in a multiple source, multiple destination network will generally impose a substantial processing burden on the system. Generally, each source-destination pair will have its own set of sequence numbers so as to uniquely identify each source and destination within the system. The total number of sequence number sets will generally increase exponentially with increasing network size. For example, where ten processors communicate with ten memory boards, one hundred sequence number sets may be required. Whereas, for one thousand processors and one thousand memory boards, one million sequence number sets will generally be required in order to uniquely identify each source-destination pair. Such a large number of sequence number sets generally increases the burden on each source and destination by increasing the number of bits required to identify the message being transmitted.

Furthermore, such a system is difficult to effectively scale down. For example, where a system is initially designed to accommodate one million sequence number sets, the system may handle the overhead associated with the sequence numbers in an acceptable manner. However, if the same chips were to be used in a small topology involving, for example, one processor and one memory board, a chip set designed to handle one million possible sequence number sets will be employed to handle just one sequence number set for the small topology, thereby imposing an enormous and unproductive overhead cost to operation of the small topology.

Generally, a typical action responsive to a determination by a destination device that a transmission is faulty, or was lost, is to issue a request to the source to retransmit the pertinent information. In a single source, single destination system, such an approach is generally quite feasible. However, in a multiple source, multiple destination system, the very data which identifies the source of the data could be lost during transmission, making it extremely difficult for the destination device to issue a retransmission request. Accordingly, the use of sequence numbers to identify transactions in a multiple source, multiple destination system may make both detection and correction of a transmission error very difficult.

One alternative to the use of a distinct set of sequence numbers for every source-destination pair in a large network is the use of software to dynamically allocate channels at a processor or other communication component in the network. Generally, software operates to allocate a channel within a processor to communicate with another component in source-destination pair for a period of time. When communication is established with a different component, the channel may be re-allocated for communication with the new component and a set of sequence numbers established on-the-fly. This approach generally avoids the requirement of having an immense number of sets of sequence numbers but generally imposes latency on the communication process. Specifically, the time required to initially allocate and re-allocate channels for communication with different components or devices requires a finite period of time. In networks where communication between devices generally experiences very short latencies, such as connections between chips in a computer system, the time required for channel allocation and re-allocation generally imposes an unacceptable delay.

Another alternative to the option of using distinct sequence number sets for all source-destination device pairs in a network is the use of pre-allocated hardware channels on all participating communication devices. FIGURE 2 depicts destination devices 202 and 203 having dedicated channel connections to processor agent device 201. This approach presents the advantage of not incurring the communication latencies generally caused by the previously discussed software-based dynamic channel allocation approach, but imposes a substantial requirement for additional hardware functionality on each participating communication device. This added hardware requirement will generally impose substantial additional system cost.

The use of sequence numbers to identify transactions generally imposes a forced ordering of transactions between a source and a destination even where such ordering is not required by higher protocol levels. This is because the destination generally only identifies missed packets by receiving a series of packets whose sequence numbers are not in proper order. However, in a multiple source, multiple destination network, packets with higher

sequence numbers may arrive at a destination device prior to an earlier sent packet with a lower sequence number due to the higher numbered packet having taken a more expeditious path through the network. This out-of-order arrival does not necessarily indicate that a packet has been lost. Therefore, a buffer will generally be employed to hold a packet with a higher sequence number so that packets with lower sequence numbers are allowed to pass through the buffer and arrive at the destination first. Such forced ordering thereby generally operates to slow down the transmission of transactions and thereby the network as a whole.

Therefore, it is a problem in the art that the use of distinct sequence number sets for all source-destination pairs in a multiple source, multiple destination network (or large complex network) imposes a burden on communication and computation in the network.

It is a further problem in the art that the use of sequence numbers in a multiple source, multiple destination network may prevent accurate identification of the device which transmitted a data packet identified as defective at a destination device.

It is a still further problem in the art that the use of sequence numbers for packet identification in large complex network generally imposes a need for forced ordering of transactions in said network.

It is a still further problem in the art that the use of software based dynamically allocated channels for sequence numbering of data packets generally imposes unacceptable latency on communication within a computer network.

It is a still further problem in the art that the use of dedicated pre-allocated hardware channels for sequence numbering of data packets in a network generally imposes a substantial cost burden on the system.

SUMMARY OF THE INVENTION

These and other objects, features and technical advantages are achieved by a system and method which enables a storage system employing temporary and reusable entries for monitoring transactions between data sources and destinations across an unordered unreliable transport mechanism. The monitoring of a transaction, and the status thereof, may be performed at a source or destination site. Monitoring of transactions at intermediate points within the transport mechanism is preferably unnecessary with the present invention. Where repetition of the communication of transaction packets or messages is necessary, end to end retry is preferably employed. Generally, end to end retry refers to initiation of a retransmission from one end of a communication link rather than attempting such retransmission from an intermediate point within this link.

In a preferred embodiment, a sequence of transactions is transmitted between a source site, and a destination site, to manage the delivery and acknowledgment of packets between the two sites and to provide unambiguous information regarding the status of all transactions initiated by the source. The inventive mechanism preferably provides exactly once semantics between source and destination over an unordered unreliable network.

In a preferred embodiment, a database, which may be a content addressable memory, associated with each destination device operates as a registry for pending transactions within the network. A preferred embodiment approach involves transmitting a sequence of transaction messages between the source and destination to ensure proper execution of the transaction and to maintain information on the status of the transaction. An exemplary sequence of transaction messages preferably includes a reservation request, a reservation response, a release request, and a release response, which transaction messages are discussed in greater detail in the Detailed Description.

In a preferred embodiment, the transport mechanism is a computer network linking together a plurality of processors, switching devices, memory devices, and other data devices in an unordered and unreliable network. Alternatively, the transport mechanism may include any set of connections between data storage or data processing devices including but not

limited to: local area networks and wide area networks. Herein, “unreliable” refers to a condition wherein a source initiating a transaction may not depend on the pertinent network to transmit the message to the desired destination, and may not assume that the destination will execute the desired transaction.

5 Instead, verification of both the transmission of a request by a source and of the successful completion of a requested transaction by a destination is preferred for the inventive mechanism to conclude that a transaction has been successfully completed. The “end to end” reporting and verification system contemplated herein preferably avoids having to perform enquiries at intermediate points in a network linking the source and destination and removes a need to assume that communication along this network has been successful.

10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

20

25

Herein, a site which initiates a sequence of transaction messages to effect a transaction at a another site is a “source site,” and the “other” site at which the transaction is executed is a “destination site.” The source and destination sites may each include a single component or may include a plurality of components. The terms “source” and “destination” are generally used synonymously with “source site” and “destination site,” respectively. In the Detailed Description, the source agent is a device specifically associated with a processor at a source site and the “destination memory control” is a device specifically associated with a memory board at a destination site. However, the term “agent device” is used herein to refer generally to a component or device which serves as an agent to a another device at that site, such as a processor or memory board. Therefore, the term agent device preferably applies to both the “source agent” and the “destination memory control.” Since, both the processor and memory board are devices which are being served by other devices, they are referred to generically herein as “master devices.” Although the term device generally refers to a single hardware component, the various devices at a site need not be discrete components, but may be integrated into a single package. Furthermore, the function of such devices may optionally be performed in software.

Herein, the term “transaction” includes an operation to be performed at a site remote from source site requesting the operation as well as the communication steps necessary to

effect this operation. An operation may be a memory read, memory write, device control operation, or other data transfer or manipulation operation. A sequence of communication steps is preferably undertaken to complete a transaction, and the information packages transmitted in connection with a transaction are referred to herein as "transaction packets," or "transaction messages." These terms may be further qualified according to the content of each message. For brevity, a transaction which includes a read operation is referred to as a "read transaction," and a transaction which includes a write operation is referred to as a "write transaction."

10 Herein, the term "exactly once semantics" generally refers to a condition wherein each transaction in the system is preferably acted upon only once. Generally, an error condition exists if a transaction is either not performed at all (i.e. zero occurrences), or performed two or more times.

In a preferred embodiment, a source may be a computer processor having an associated agent chip for conducting communication regarding the status of transactions, thereby removing the processor from a critical path for such communication. In a preferred embodiment, the destination is a memory board or other data storage device which preferably has an agent device for handling communication pertaining to transactions with other devices. One possible type of agent device for the data storage device is a destination memory control. Preferably, both the source and destination sites include databases for preserving information 20 regarding the status of a transaction while said transaction is active within the network.

25 In a preferred embodiment, agent devices included in source and destination sites operate to ensure delivery of transactions across an unreliable network. The use of agent devices for such a maintenance function preferably improves the efficiency of the source and destination devices by relieving them of the processing burden of monitoring transaction communication. Preferably, databases are included at the source and destination sites for storing information identifying the transactions active in the network and the status of these transactions. Preferably, data locations in these databases are used temporarily for particular transactions and are infinitely reusable. Preferably, when a new transaction is generated, an

entry associated with that transaction is stored in one or more databases and continues to be used to reference that transaction until the transaction is no longer active in the network. Upon completion of a transaction, the entry for that transaction is preferably released for reuse in association with a new transaction. The data entry mechanism of the present

5 invention is thereby preferably able to serve an infinite stream of successive transactions with a finite number of database entries by releasing device entries upon conclusion of a transaction, and reallocating them upon initiation of a new transaction. Where a transaction, for any reason, cannot be completed, a response packet containing relevant error information is preferably provided to any device awaiting a response in connection with the incomplete transaction. If no device is waiting for a response in connection with an incomplete transaction, error information associated with the transaction is preferably logged locally before resources associated with the incomplete transaction are enabled to be used for other purposes.

10 11 12 13 14 15 16 17 18 19

20

In a preferred embodiment, transaction messages may be transmitted by the source sites in any order and received by the destination sites in any order with no adverse effect on the ability of the inventive mechanism to record and manage communication activities between sources and destinations of transactions in a transport mechanism, which may be a computer network. This feature preferably provides for transaction ordering flexibility, thereby removing a constraint on the network generally present when employing sequence numbers. This represents an advantage over the prior art where rigid ordering of transactions generally operates as a constraint on network activity, and in certain cases, causes otherwise executable transactions to remain idle, to enable transactions with lower sequence numbers to be processed first.

25

In a preferred embodiment of the present invention, the delays associated with allocating a channel for sequence numbers in software at a source or destination site is preferably unnecessary because of the operation of the agent devices and databases at each site as discussed above. Furthermore, the use of temporary reusable entries within the databases for managing the communication of transactions over the network preferably

renders the allocation of dedicated hardware channels within a communication device unnecessary.

In a preferred embodiment, if the source does not receive a reply packet responsive to a packet transmitted by the source in a pre-determined time-out period, the source preferably retransmits its packet. This retransmission presents the possibility of having two identical source-issued messages propagating through the system at the same time. In order to prevent this from occurring, the source includes a time stamp in the outgoing packet and establishes a time-out threshold for the outgoing packet, which is preferably shorter than the time-out threshold for the transmitting source, thereby enabling any device in the network to drop the packet from the network when the packet time-out threshold is reached. The system thereby preferably ensures that there will never be two packets being simultaneously transmitted through the network.

Therefore, it is an advantage of a preferred embodiment of the present invention that the computationally burdensome operation of designating distinct sequence number sets for source-destination pair in a communication network is unnecessary.

It is a further advantage of a preferred embodiment of the present invention that forced ordering of transactions is not required, thereby enabling transactions to be efficiently processed at a destination device in the order in which they are received.

It is a still further advantage of a preferred embodiment of the present invention that pre-allocated hardware channels are not required in order to manage the communication of transactions.

It is a still further advantage of a preferred embodiment of the present invention that greater certainty regarding the status of a transaction in the network is provided than was available using the sequence number identification system of the prior art.

The foregoing has outlined rather broadly the features and technical advantages of the present invention in order that the detailed description of the invention that follows may be better understood. Additional features and advantages of the invention will be described

hereinafter which form the subject of the claims of the invention. It should be appreciated by those skilled in the art that the conception and specific embodiment disclosed may be readily utilized as a basis for modifying or designing other structures for carrying out the same purposes of the present invention. It should also be realized by those skilled in the art that such equivalent constructions do not depart from the spirit and scope of the invention as set forth in the appended claims.

BRIEF DESCRIPTION OF THE DRAWING

For a more complete understanding of the present invention, and the advantages thereof, reference is now made to the following descriptions taken in conjunction with the accompanying drawing, in which:

5 FIGURE 1 depicts a multiple source, multiple destination network;

FIGURE 2 depicts an agent device and two separate destination devices wherein dedicated channels link each of the destination devices to the agent device;

10 FIGURE 3 depicts a transaction scheme employing a destination memory control device according to a preferred embodiment of the present invention;

FIGURE 4 depicts a transmission of a first of four transaction messages according to a preferred embodiment of the present invention;

15 FIGURE 5 depicts a transmission of a second of four transaction messages according to a preferred embodiment of the present invention;

FIGURE 6 depicts a transmission of a third of four transaction messages according to a preferred embodiment of the present invention;

FIGURE 7 depicts a transmission of a fourth of four transaction messages according to a preferred embodiment of the present invention; and

20 FIGURE 8 depicts a network having a plurality of source devices and destination devices the communication between which is routed over an unordered, unreliable network according to a preferred embodiment of the present invention.

DETAILED DESCRIPTION

5

FIGURE 3 depicts a transaction scheme employing a destination memory control device 305 according to a preferred embodiment of the present invention. In a preferred embodiment, source agent 302 is associated with processor 301. Requirements database 303 preferably cooperates with source agent 302 to store the status of transactions and transaction data packets. Network 304 is generally a multiple source, multiple destination computer network but may alternatively be a single device, such as a router, interposed between a single source device and a single destination device. In another alternative embodiment, network 304 may be a local area network or wide area network.

10
11
12
13
14
15
16
17
18
19

In a preferred embodiment, Content Addressable Memory (CAM) 306 is a memory device which allows any one of its entries to be written to by an external device. CAM 306 preferably includes a feature wherein, when presented with the identification of a source device, or other external device, and an identification of a transaction of interest, memory control 305 will compare the presented transaction identification to all current valid entries in the CAM 306 to determine whether the pertinent transaction is present in the CAM. CAM 306 preferably operates as a database for storing information regarding currently active transactions in network 300.

20

In a preferred embodiment, processor 301 operates to conduct a transaction with memory board 307, which transaction may be a memory read or memory write operation. The source agent 302 and database 303 at the source site and memory control 305 and CAM 306 at the destination site preferably operate to provide management of the transaction and preservation of data on the progress and status of the transaction so as to minimize latency, avoid duplication of transaction steps, and enable system 300 to possess unambiguous information regarding the status of all transactions pending in system 300 at any given time.

25

Alternatively, the source 302 and destination 305 devices need not be a processor and memory, respectively, but may include a range of data storage and manipulation devices. For example, a processor to processor transaction, involving a read, write, or other operation is suitable for application of the inventive mechanism. In another alternative embodiment, both

memory boards and a processors may be either sources or destinations in the configuration of FIGURE 3.

In a preferred embodiment, a sequence of four transaction messages or packets are communicated between source agent 302 and memory control 305 in order to establish a status for each transaction at appropriate points within system 300. It will be appreciated that an alternate number of transactions less than or more than four may be employed to perform the functions of the inventive mechanism, and all such variations are within the scope of the present invention. Preferably, the protocol of the present invention ensures that although individual transaction packets may be issued or received more than once, the transaction being implemented by the various transaction packets is performed only once.

FIGURE 4 depicts a transmission of a first of four transaction messages according to a preferred embodiment of the present invention. Reference is made to both FIGURES 3 and 4 in the following discussion. Source agent 302 preferably sends reservation request 307 via network 304 to memory control 305. Reservation request 307 preferably includes information describing a transaction being requested by source agent 302 on behalf of processor chip 301 as well a transaction identification to identify the transaction to memory control 305. Reservation request 307 preferably further includes a time stamp 402 to identify to any device within network 304 encountered by request 307 the time at which reservation request 307 was transmitted by agent 302. The information incorporated in time stamp 402 preferably enables reservation request 402 to be dropped from the system, by any device within the system, upon expiration of a predetermined period (the packet time-out period) beginning with the time indicated in time stamp 402.

In a preferred embodiment, agent 302 includes a timing function 401 for recording the time elapsed since the transmission of reservation request 307. Preferably, agent 302 records, in database 303, the fact and time of transmission of reservation request 307. The transmission of reservation request 307 initiates timer 401 in agent 302 to determine whether reservation response 308 is received within a predetermined period of transmission of reservation request 307. If the source agent time-out period elapses without receipt of

reservation response 308 at agent 302, agent 302 retransmits reservation request 307. The time-out period for dropping reservation request 307 from the system is preferably smaller than the time period established in agent 302 for retransmitting the reservation request, thereby ensuring that any prior reservation request packets have been dropped from the 5 system prior to transmission of a duplicate reservation request. In this manner, the inventive mechanism preferably operates to ensure that there will more than one reservation request pertaining to the same transaction traveling through network 304 at the same time.

In a preferred embodiment, when destination memory control 305 receives reservation request 307, memory control 305 preferably determines whether there is a pre-existing entry in CAM 306 corresponding to reservation request 307. If no corresponding entry is found in CAM 306, memory control 305 preferably concludes that reservation request 307 is being received at memory control 305 for the first time. In this case, the transaction associated with reservation request 307 will preferably be acted upon by writing data to, or reading data from, memory 307.

In a preferred embodiment, if an entry is found in CAM 306 which corresponds to reservation request 307, memory control 306 preferably concludes that reservation request 307 is the result of a resend operation. Generally, at least two options are available to memory control 305 in this case. Where the reservation response generated in response to the packet of which reservation request 307 is a duplicate, is completely stored in CAM 306, 20 control 305 retransmits the stored reservation response as reservation response 308 responsive to reservation request 307. Where, for example, the pertinent transaction is a read operation and data from the read operation is not stored in CAM 306, control 305 will acquire the data from memory board 307 corresponding to the pertinent read operation, and transmit this data as part of reservation response 308 to source 302.

25 In a preferred embodiment, data resulting from a read operation is not automatically stored in CAM 306. However, where a read transaction is still active in the network, and a memory location which is read from in that read operation will be erased by a prospective write operation or other computing procedure, control 305 preferably loads data from the

memory location (which is soon to be over-written) into the CAM so that it will be available in the event that one or more subsequent reservation requests associated with the pertinent read transaction are received by control 305.

FIGURE 5 depicts a transmission of a second of four transaction messages according to a preferred embodiment of the present invention. Reference is made to both FIGURES 3 and 5 in the following discussion. FIGURE 5 depicts the transmission of reservation response 308 by memory control 305 via network 304 to agent 302. Preferably, generating a reservation response 308 responsive to reservation request 307 includes the steps of generating the data associated with the response, combining this response data with an appropriate transaction identifier to form a reservation response packet, and finally, transmitting reservation response 308.

In a preferred embodiment, the reservation response data is transaction dependent. In the case of a memory “read” operation, the reservation response data preferably comprises the data acquired from memory or memory board 307 as a result of a “read” transaction included in reservation request 307. In the case of a memory “write” operation, there is generally no data to return to source agent 302. The reservation response responsive to a write operation will therefore preferably include an acknowledgment that a requested write operation was successful. Once the reservation response has been prepared, it is stored in the CAM 306 and preferably remains there until a release request packet is received by the memory control 305.

In a preferred embodiment, where a reservation response responsive to the reservation request is not present in the CAM, reservation response 308 is generated and transmitted to agent 302 via network 304 as previously discussed. Where a reservation response corresponding to the reservation request is present in the CAM, indicating that a duplicate reservation request has been received, the already stored reservation response is preferably transmitted to agent 302. Preferably, minimal computational effort is required to transmit this already existing reservation response.

After memory control 305 either discovers reservation response 308 in the CAM 306 or generates the information required therefor, memory control 305 preferably transmits

reservation response to source agent 302 as discussed in connection with FIGURE 5. For the sake of this discussion, it will be assumed that source agent 302 receives reservation response intact and prior to expiration of timer 401 in source agent 302. Timer 401 preferably counts the time elapsed since the transmission of the reservation request 307 to which reservation response 308 is responsive. The system response to reservation request packets and reservation response packets which have been dropped by the system is discussed later in this application.

5 In a preferred embodiment, receipt of reservation response 308 by agent 302 preferably informs agent 302 of the status of the transaction which was included in the reservation request originally transmitted by agent 302. Preferably, reservation response 308 will indicate that the transaction read or write operation was successfully executed.

10 Alternatively however, reservation response 308 may include information indicating that a memory read or write operation was not successful. Further, memory control 305 may optionally insert into reservation response 308, details pertaining to an unsuccessful transaction and available options for correcting the problem. In a preferred embodiment, reservation response 308 returns data to source agent 302 and eventually to processor 302 resulting from execution of a successful transaction.

20 In a preferred embodiment, source agent 302, having received reservation response 308, preferably operates to inform memory control 305 that the data stored in CAM 306 pertaining to the transaction associated with reservation request 307 may be deleted so as to release space in the CAM for use with a new reservation request. It will be appreciated that the released CAM space may be used for a reservation request emerging from any processor connected to network 304 and is not limited to storing such requests only from source agent 302.

25 FIGURE 6 depicts a transmission of a third of four transaction messages according to a preferred embodiment of the present invention. Reference is made to both FIGURES 3 and 6 in the following discussion. FIGURE 6 depicts the transmission of release request 309 from source agent 302 across network 304.

In a preferred embodiment, the transmission of release request by agent 302 signifies that agent 302 and processor 301 regard the transaction transmitted by reservation request 307 as having been concluded, and that stored information pertaining to the completed transaction at a destination site, which in this case includes CAM memory 306, is no longer needed. Accordingly, upon receipt of release request 309 by memory control 305, memory control 305 preferably clears the entry in CAM 306 pertaining to the pertinent transaction.

In a preferred embodiment, as was the case with the transmission of reservation request 307, timer 401 in source agent 302 operates to count elapsed time from the transmission of release request 309. If release response 310 is received before timer 401 reaches a retransmission threshold point, the timer is preferably reset and no retransmission occurs. However, if the timer threshold point is reached prior to receipt of release response 310 at source agent 302, source agent 302 preferably retransmits release request 309. Without any other action, a retransmission operation might leave open the possibility of having two identical release requests in transmission through the system at the same time.

Therefore, in a preferred embodiment, a timing mechanism is implemented in the system to cause a release request to “time out” and drop out of the system after a predetermined period has expired since transmission of that release request. As was described in connection with the reservation request, a time stamp 601 is preferably included in release request 309 which may be read by a number of devices which may be encountered by the release request as it propagates through network 304. Preferably, any one of these devices is enabled to read to keep current time and remove release request 309 from the system if the release request packet time-out period has expired.

In a preferred embodiment, in order to ensure that the release request “times out” before source agent 302 issues a retransmission request, the time-out threshold at which system devices drop the release request from the system is preferably shorter than the time-out threshold which source agent 302 employs to signal retransmission of the release request.

FIGURE 7 depicts a transmission of a fourth of four transaction messages according to a preferred embodiment of the present invention. Reference is made to both FIGURES 3

and 7 in the following discussion. Preferably, memory control 305 transmits release response 310 across network 304 to agent 302.

In a preferred embodiment, transmission of release response 310 by memory control 305 indicates that the transaction which is the subject of the preferably four transaction 5 messages has been completed and that the CAM entry for this transaction has been released, thereby freeing up space in the CAM 306 for a new reservation request. Preferably, upon receipt of release response 310, source agent 302 appropriately concludes that the transaction has been successfully concluded and proceeds to delete an entry corresponding to the completed transaction in database 303.

In a preferred embodiment, source agent 302 receives the desired results of a transaction when reservation response 308 is delivered to source agent 302. Source agent 302 then preferably communicates this information in a timely manner to processor 301. Subsequently, the third and fourth transaction messages, the release request and release response, respectively, constitute follow-up operations to appropriately inform memory controller 305 and source agent 302 of the status of the transaction at their respective locations and to free up temporarily allocated storage space in database 303 (on the source agent end) and in CAM 306 (on the destination memory control end). Employing source agent 302 for the management of transaction messages and entering and deleting entries in database 303 preferably beneficially enables processor 301 to concentrate on performing 20 operations on the received data, instead of being burdened with communication and database management tasks.

In a preferred embodiment, the provision of services by devices other than the processor, including deleting database and CAM entries for completed transactions and the related communication in both directions of status information, increases processor efficiency 25 and reduces processing and communication latency. The improvement (reduction) in processing latency provided by delegating tasks to the agent 302 and memory control 305 would generally not have been available employing the sequence number systems of the prior art, where processing time would have been spent allocating channels and establishing

sequence numbering systems on both ends of a source-destination pair. The inventive approach therefore provides a significant latency advantage over the systems of the prior art.

Generally, packets transmitted across an unreliable network, such as network 304 may be dropped or corrupted. In a preferred embodiment, the problem of network unreliability and the attendant possibility of packets being dropped is preferably handled by employing time-outs at the source of both reservation requests and release requests as well as inserting a time stamp in both of the reservation request packet 307 and release request packet 309. The interaction time-outs at source agent 302 and time-outs occurring at dispersed locations within the system employing the time stamp in the packets has already been discussed in detail and will therefore not be repeated in detail in this section. While the timer in source agent 302 triggers retransmission of a packet, the combination of the time stamp and the ability of dispersed devices within network 304 to check the current time against a time stamp and time-out period included in a packet, operates to drop out of the system, packets which have "timed out." Preferably, the time-out threshold of the packets traveling through the system is set to be lower than the time out employed by the agent 302 to trigger retransmission of a previously sent packet.

In a preferred embodiment, source agent 302 employs a timer to determine when to retransmit a message where an expected response has not been received. Source agent 302 also includes a time stamp in both the reservation request 307 and release request 309 in order to enable the system to drop these requests from the system if they have not reached the destination memory control 305 within an pre-determined time. In a preferred embodiment, memory control 305 does not perform a similar timing operation for messages transmitted to agent 302. Furthermore, memory control 305 preferably does not include time stamps in reservation response 308 and release response 310. It will be appreciated however, that in an alternative embodiment, the process of counting the time elapsed since a message was transmitted could be performed by memory control 305 alone, or by both source agent 302 and memory control 305. Furthermore, in an alternative embodiment, memory control 305

could include time stamps in the packets for reservation response 308 and release response 310, and all such variations are included within the scope of the present invention.

In a preferred embodiment, agent 302 retransmits packets when its logic indicates that a packet responsive to the transmitted packet should have, but has not been, received after a predefined time period has elapsed since a prior transmission of the packet outbound from the source or agent. Source agent 302 preferably retransmits the relevant packet assuming that the previously transmitted packet has been dropped from the system. Generally, the effect on the system of a dropped packet depends upon which of the four transaction communication packets was dropped.

In a preferred embodiment, if a reservation request is dropped en route to memory control 305, agent 302 will time out according to its schedule, and preferably retransmit a duplicate reservation request by accessing database 303. It will be appreciated that agent 302 is not limited to sending the reservation request only twice, but rather, may repeat the transmission as many times as is necessary to ensure completion of this first transaction packet transmission. However, while transient problems in system 300 may account for the loss of one or two packets, agent 302 may establish a threshold number of failed transmissions above which it regards the network as containing a permanent failure and works to arrange for an alternative method of accomplishing the transmission of the required packet.

In a preferred embodiment, where memory control 305 receives a reservation request representing a transaction not found in the CAM 306, control 305 treats the incoming packet as a first reservation request transmission, preferably establishes a new entry in CAM 306, and preferably produces a reservation response responsive to the reservation request packet. Where memory control 305 receives a reservation request representing a transaction which is found in CAM 306, memory control 305 preferably concludes that a reservation response corresponding to the currently requested transaction was produced and transmitted, but dropped from the system. Accordingly, it is preferably not necessary to generate a new reservation response in response to the duplicate reservation request. Preferably, memory

control 305 generates a duplicate reservation response with information stored by the original reservation response. In a preferred embodiment, the above logic applies in similar manner to the transmission by source agent 302 of release request 309 and the packet responsive thereto, release response 310. In this case, the logic is generally reversed. Since release
5 request 309 generally causes memory control 305 to delete the entry in the CAM 306 for the pertinent transaction, upon receipt of a first release request 309 without any dropped packets, there should not be an entry corresponding to the transaction associated with the release request.

Accordingly, in a preferred embodiment, if a release request 309 is received which does have an entry in the CAM 306, memory control 305 preferably considers the release request 309 to be a first transmission, appropriately acts to delete the pertinent entry in CAM 306, and generates a release response 310. If a release request packet was dropped before reaching memory control 305, memory control 305 preferably interprets the release request packet to be a first transmission, as described above.

Where a release request is received which does not match an entry in CAM 306, memory control preferably considers release request 309 to be a duplicate transmission, and generates and transmits release response 310 to agent 302 across network 304. It will be appreciated that where the release request does not match any entry in CAM 306, there is no entry in CAM 306 which is suitable for deletion. Memory control 305 then preferably generates and transmits a release response 310 to agent 302 across network 304. The sequence of events described above is consistent with a prior release response 310 having been dropped before reaching agent 302.

For purposes of further discussion, the processor and memory are generally referred to herein as "master" devices, and the source agent and destination memory control as "servant" devices. It will be appreciated that the inventive principles are not limited to the master and servant devices being hardware components. The master devices could alternatively be software programs which delegate selected tasks to "servant" programs, or subroutines, so as to remove the delegated tasks from their respective critical paths and thereby execute more
25

efficiently. Likewise, the servant devices may be programs or sub-routines which serve main programs and operate to perform tasks which relieve their respective main programs of certain time consuming tasks. Preferably, the master and servant software components would implement a similar protocol for preserving transaction results and storing transaction packet message statuses as is described above in connection with hardware versions of the master and servant devices.

FIGURE 8 depicts a network having a plurality of source and destination devices, the communication between which is routed over an unordered, unreliable network according to a preferred embodiment of the present invention. FIGURE 8 depicts agent 803 associated with processor 801 and agent 804 associated with processor 802. The sequence of dots in between agents 803 and 804 is intended to indicate the presence of a plurality of additional agents also having associated processors. Similarly, memory controls 805 and 806 are associated, respectively, with memory boards 807 and 808. The various agents and memory controls are generally all connected to network 304. In the following paragraph, the combination of an agent and a processor, when acting to request a transaction from a remote location on the network 800, is generally referred to as a source site, and the combination of a memory control and a memory board when operating to service a transaction request received from a remote source site, is generally referred to as a destination site.

In a preferred embodiment, a transaction requested by a source, executed at a destination, and implemented via a sequence of preferably four transaction packets or messages, as described in connection with FIGURES 3-7 may generally be accomplished between any source-destination combination connected across network 800. For example, where there are one hundred source sites and one hundred destination sites, generally ten thousand different source-destination pairs are available. Although one embodiment of the source-destination pairing may include a source which includes a processor and a destination which includes a memory board, it will be appreciated that source sites are not limited to having processors, and that destination sites are not limited to having memories or memory boards. In alternative embodiments, both processors along with associated equipment, and

10
15
20
25

memory boards along associated equipment, may operate as either source sites or destination sites, and all such variations are included within the scope of the present invention.

Although the present invention and its advantages have been described in detail, it should be understood that various changes, substitutions and alterations can be made herein without departing from the spirit and scope of the invention as defined by the appended claims. Moreover, the scope of the present application is not intended to be limited to the particular embodiments of the process, machine, manufacture, composition of matter, means, methods and steps described in the specification. As one of ordinary skill in the art will readily appreciate from the disclosure of the present invention, processes, machines, manufacture, compositions of matter, means, methods, or steps, presently existing or later to be developed that perform substantially the same function or achieve substantially the same result as the corresponding embodiments described herein may be utilized according to the present invention. Accordingly, the appended claims are intended to include within their scope such processes, machines, manufacture, compositions of matter, means, methods, or steps.

WHAT IS CLAIMED IS:

1. A method for executing a transaction in a network having a source site and a destination site, the method comprising the steps of:

5 transmitting an initial transaction request message from said source site to said destination site;

receiving said transaction request message at said destination site;

establishing an association of a data entry with said transaction in a destination database; and

10 preserving said association of said data entry with said transaction in said destination database so long as said transaction is active in said network.

2. The method of claim 1, comprising the further step of:

executing said transaction at said destination site, thereby producing transaction results.

3. The method of claim 2, comprising the further step of:

where a prospective operation will override said transaction results in a memory board,

5 storing said transaction results in said destination database, thereby enabling retransmission of said transaction results if a further reservation request message is received at said destination site.

4. The method of claim 2, comprising the further step of:

transmitting said transaction results to said source site over said network.

5. The method of claim 1 comprising the further step of:

transmitting another transaction request message if no response is received from said destination site at said source site within a source site time-out period.

6. The method of claim 5, comprising the further step of:

deleting said initial transaction request message from the network if said transaction request message does not reach said destination site within a request message time-out period, wherein said source site time-out period exceeds said request message time-out period to prevent having two transaction request messages simultaneously in transmission through said network.

7. The method of claim 4, comprising the further steps of:

upon receiving a duplicate transaction request message,

identifying the data entry in the destination database established for said transaction;

acquiring said transaction results; and

retransmitting said acquired transaction results to said source site.

8. The method of claim 7, wherein the step of acquiring comprises:

retrieving said transaction results from said destination database.

9. The method of claim 7, wherein the step of acquiring comprises:

executing said transaction in response to said duplicate transaction request message, thereby producing said transaction results.

10. The method of step 4, comprising the further steps of:
receiving said transmitted transaction results at said source site; and
transmitting, from said source site to said destination site, a release request to delete
said data entry associated with said transaction in said destination database.

11. The method of step 10, comprising the further steps of:
receiving at said destination site, said release request to delete said data entry
associated with said transaction; and
deleting, within said destination database, said data entry associated with said
transaction, thereby liberating space in said destination database.

12. The method of step 11, comprising the further step of:
transmitting, from said destination site to said source site, a release response message,
thereby indicating that said data entry associated with said transaction in said destination
database has been deleted.

13. The method of step 1, comprising the further step of:
wherein the source site includes a processor and an agent device,
delegating said step of transmitting said initial transaction request message to said
agent device.

14 A system for reliably executing a transaction at a destination site requested by a source site, the system comprising:

means for transmitting an initial transaction request message to said destination site from said source site;

5 means for executing a transaction associated with said initial transaction request message at said destination site;

a reservation database at said destination site for storing information uniquely identifying said transaction.

15. The system of claim 14, wherein the reservation database is a content addressable memory.

16. The system of claim 14, wherein the source site comprises:

a processor; and the destination site comprises:

a memory.

17. The system of claim 16, wherein the source site further comprises:

a processor agent device for conducting communication with said destination site, thereby enabling said processor to efficiently concentrate on other tasks.

18. The system of claim 17, wherein the source site further comprises:

a source site database for preserving an identification and a status of said transaction until said transaction is complete.

19. The system of claim 16, wherein the processor agent device comprises:
a timer for initiating a retransmission of said transaction request message if no
message responsive to said initial transaction request message is received at said processor
agent device upon expiration of a retransmission time-out period.

20 A system for executing a transaction in a network having a source site and a destination site, the system comprising:

means for transmitting an initial transaction request message from said source site to said destination site;

5 means for receiving said transaction request message at said destination site;

means for establishing an association of a data entry with said transaction in a destination database; and

means for preserving said association of said data entry with said transaction in said destination database so long as said transaction is active in said network.

10

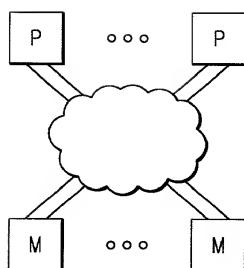
**PROTOCOL FOR INSURING EXACTLY ONCE SEMANTICS OF
TRANSACTIONS ACROSS AND UNORDERED, UNRELIABLE NETWORK**

ABSTRACT OF THE DISCLOSURE

The present invention implements a mechanism for reliably communicating transaction messages between source and destination devices connected across an unreliable network, where a transaction is an operation requested by a source device of a destination device and where a plurality of messages are exchanged between source and destination in order to ensure completion of the transaction. Because the network linking the source and destination is unreliable, the source and destination operate to preserve data regarding the transaction messages which they have received and transmitted. If responsive messages are not timely received, they are generally resent. A dual timing system is preferably implemented to ensure that there are never two identical messages simultaneously in transmission through the network.

5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
140
145
150
155
160
165
170
175
180
185
190
195
200
205
210
215
220
225
230
235
240
245
250
255
260
265
270
275
280
285
290
295
300
305
310
315
320
325
330
335
340
345
350
355
360
365
370
375
380
385
390
395
400
405
410
415
420
425
430
435
440
445
450
455
460
465
470
475
480
485
490
495
500
505
510
515
520
525
530
535
540
545
550
555
560
565
570
575
580
585
590
595
600
605
610
615
620
625
630
635
640
645
650
655
660
665
670
675
680
685
690
695
700
705
710
715
720
725
730
735
740
745
750
755
760
765
770
775
780
785
790
795
800
805
810
815
820
825
830
835
840
845
850
855
860
865
870
875
880
885
890
895
900
905
910
915
920
925
930
935
940
945
950
955
960
965
970
975
980
985
990
995
1000
1005
1010
1015
1020
1025
1030
1035
1040
1045
1050
1055
1060
1065
1070
1075
1080
1085
1090
1095
1100
1105
1110
1115
1120
1125
1130
1135
1140
1145
1150
1155
1160
1165
1170
1175
1180
1185
1190
1195
1200
1205
1210
1215
1220
1225
1230
1235
1240
1245
1250
1255
1260
1265
1270
1275
1280
1285
1290
1295
1300
1305
1310
1315
1320
1325
1330
1335
1340
1345
1350
1355
1360
1365
1370
1375
1380
1385
1390
1395
1400
1405
1410
1415
1420
1425
1430
1435
1440
1445
1450
1455
1460
1465
1470
1475
1480
1485
1490
1495
1500
1505
1510
1515
1520
1525
1530
1535
1540
1545
1550
1555
1560
1565
1570
1575
1580
1585
1590
1595
1600
1605
1610
1615
1620
1625
1630
1635
1640
1645
1650
1655
1660
1665
1670
1675
1680
1685
1690
1695
1700
1705
1710
1715
1720
1725
1730
1735
1740
1745
1750
1755
1760
1765
1770
1775
1780
1785
1790
1795
1800
1805
1810
1815
1820
1825
1830
1835
1840
1845
1850
1855
1860
1865
1870
1875
1880
1885
1890
1895
1900
1905
1910
1915
1920
1925
1930
1935
1940
1945
1950
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
2045
2050
2055
2060
2065
2070
2075
2080
2085
2090
2095
2100
2105
2110
2115
2120
2125
2130
2135
2140
2145
2150
2155
2160
2165
2170
2175
2180
2185
2190
2195
2200
2205
2210
2215
2220
2225
2230
2235
2240
2245
2250
2255
2260
2265
2270
2275
2280
2285
2290
2295
2300
2305
2310
2315
2320
2325
2330
2335
2340
2345
2350
2355
2360
2365
2370
2375
2380
2385
2390
2395
2400
2405
2410
2415
2420
2425
2430
2435
2440
2445
2450
2455
2460
2465
2470
2475
2480
2485
2490
2495
2500
2505
2510
2515
2520
2525
2530
2535
2540
2545
2550
2555
2560
2565
2570
2575
2580
2585
2590
2595
2600
2605
2610
2615
2620
2625
2630
2635
2640
2645
2650
2655
2660
2665
2670
2675
2680
2685
2690
2695
2700
2705
2710
2715
2720
2725
2730
2735
2740
2745
2750
2755
2760
2765
2770
2775
2780
2785
2790
2795
2800
2805
2810
2815
2820
2825
2830
2835
2840
2845
2850
2855
2860
2865
2870
2875
2880
2885
2890
2895
2900
2905
2910
2915
2920
2925
2930
2935
2940
2945
2950
2955
2960
2965
2970
2975
2980
2985
2990
2995
3000
3005
3010
3015
3020
3025
3030
3035
3040
3045
3050
3055
3060
3065
3070
3075
3080
3085
3090
3095
3100
3105
3110
3115
3120
3125
3130
3135
3140
3145
3150
3155
3160
3165
3170
3175
3180
3185
3190
3195
3200
3205
3210
3215
3220
3225
3230
3235
3240
3245
3250
3255
3260
3265
3270
3275
3280
3285
3290
3295
3300
3305
3310
3315
3320
3325
3330
3335
3340
3345
3350
3355
3360
3365
3370
3375
3380
3385
3390
3395
3400
3405
3410
3415
3420
3425
3430
3435
3440
3445
3450
3455
3460
3465
3470
3475
3480
3485
3490
3495
3500
3505
3510
3515
3520
3525
3530
3535
3540
3545
3550
3555
3560
3565
3570
3575
3580
3585
3590
3595
3600
3605
3610
3615
3620
3625
3630
3635
3640
3645
3650
3655
3660
3665
3670
3675
3680
3685
3690
3695
3700
3705
3710
3715
3720
3725
3730
3735
3740
3745
3750
3755
3760
3765
3770
3775
3780
3785
3790
3795
3800
3805
3810
3815
3820
3825
3830
3835
3840
3845
3850
3855
3860
3865
3870
3875
3880
3885
3890
3895
3900
3905
3910
3915
3920
3925
3930
3935
3940
3945
3950
3955
3960
3965
3970
3975
3980
3985
3990
3995
4000
4005
4010
4015
4020
4025
4030
4035
4040
4045
4050
4055
4060
4065
4070
4075
4080
4085
4090
4095
4100
4105
4110
4115
4120
4125
4130
4135
4140
4145
4150
4155
4160
4165
4170
4175
4180
4185
4190
4195
4200
4205
4210
4215
4220
4225
4230
4235
4240
4245
4250
4255
4260
4265
4270
4275
4280
4285
4290
4295
4300
4305
4310
4315
4320
4325
4330
4335
4340
4345
4350
4355
4360
4365
4370
4375
4380
4385
4390
4395
4400
4405
4410
4415
4420
4425
4430
4435
4440
4445
4450
4455
4460
4465
4470
4475
4480
4485
4490
4495
4500
4505
4510
4515
4520
4525
4530
4535
4540
4545
4550
4555
4560
4565
4570
4575
4580
4585
4590
4595
4600
4605
4610
4615
4620
4625
4630
4635
4640
4645
4650
4655
4660
4665
4670
4675
4680
4685
4690
4695
4700
4705
4710
4715
4720
4725
4730
4735
4740
4745
4750
4755
4760
4765
4770
4775
4780
4785
4790
4795
4800
4805
4810
4815
4820
4825
4830
4835
4840
4845
4850
4855
4860
4865
4870
4875
4880
4885
4890
4895
4900
4905
4910
4915
4920
4925
4930
4935
4940
4945
4950
4955
4960
4965
4970
4975
4980
4985
4990
4995
5000
5005
5010
5015
5020
5025
5030
5035
5040
5045
5050
5055
5060
5065
5070
5075
5080
5085
5090
5095
5100
5105
5110
5115
5120
5125
5130
5135
5140
5145
5150
5155
5160
5165
5170
5175
5180
5185
5190
5195
5200
5205
5210
5215
5220
5225
5230
5235
5240
5245
5250
5255
5260
5265
5270
5275
5280
5285
5290
5295
5300
5305
5310
5315
5320
5325
5330
5335
5340
5345
5350
5355
5360
5365
5370
5375
5380
5385
5390
5395
5400
5405
5410
5415
5420
5425
5430
5435
5440
5445
5450
5455
5460
5465
5470
5475
5480
5485
5490
5495
5500
5505
5510
5515
5520
5525
5530
5535
5540
5545
5550
5555
5560
5565
5570
5575
5580
5585
5590
5595
5600
5605
5610
5615
5620
5625
5630
5635
5640
5645
5650
5655
5660
5665
5670
5675
5680
5685
5690
5695
5700
5705
5710
5715
5720
5725
5730
5735
5740
5745
5750
5755
5760
5765
5770
5775
5780
5785
5790
5795
5800
5805
5810
5815
5820
5825
5830
5835
5840
5845
5850
5855
5860
5865
5870
5875
5880
5885
5890
5895
5900
5905
5910
5915
5920
5925
5930
5935
5940
5945
5950
5955
5960
5965
5970
5975
5980
5985
5990
5995
6000
6005
6010
6015
6020
6025
6030
6035
6040
6045
6050
6055
6060
6065
6070
6075
6080
6085
6090
6095
6100
6105
6110
6115
6120
6125
6130
6135
6140
6145
6150
6155
6160
6165
6170
6175
6180
6185
6190
6195
6200
6205
6210
6215
6220
6225
6230
6235
6240
6245
6250
6255
6260
6265
6270
6275
6280
6285
6290
6295
6300
6305
6310
6315
6320
6325
6330
6335
6340
6345
6350
6355
6360
6365
6370
6375
6380
6385
6390
6395
6400
6405
6410
6415
6420
6425
6430
6435
6440
6445
6450
6455
6460
6465
6470
6475
6480
6485
6490
6495
6500
6505
6510
6515
6520
6525
6530
6535
6540
6545
6550
6555
6560
6565
6570
6575
6580
6585
6590
6595
6600
6605
6610
6615
6620
6625
6630
6635
6640
6645
6650
6655
6660
6665
6670
6675
6680
6685
6690
6695
6700
6705
6710
6715
6720
6725
6730
6735
6740
6745
6750
6755
6760
6765
6770
6775
6780
6785
6790
6795
6800
6805
6810
6815
6820
6825
6830
6835
6840
6845
6850
6855
6860
6865
6870
6875
6880
6885
6890
6895
6900
6905
6910
6915
6920
6925
6930
6935
6940
6945
6950
6955
6960
6965
6970
6975
6980
6985
6990
6995
7000
7005
7010
7015
7020
7025
7030
7035
7040
7045
7050
7055
7060
7065
7070
7075
7080
7085
7090
7095
7100
7105
7110
7115
7120
7125
7130
7135
7140
7145
7150
7155
7160
7165
7170
7175
7180
7185
7190
7195
7200
7205
7210
7215
7220
7225
7230
7235
7240
7245
7250
7255
7260
7265
7270
7275
7280
7285
7290
7295
7300
7305
7310
7315
7320
7325
7330
7335
7340
7345
7350
7355
7360
7365
7370
7375
7380
7385
7390
7395
7400
7405
7410
7415
7420
7425
7430
7435
7440
7445
7450
7455
7460
7465
7470
7475
7480
7485
7490
7495
7500
7505
7510
7515
7520
7525
7530
7535
7540
7545
7550
7555
7560
7565
7570
7575
7580
7585
7590
7595
7600
7605
7610
7615
7620
7625
7630
7635
7640
7645
7650
7655
7660
7665
7670
7675
7680
7685
7690
7695
7700
7705
7710
7715
7720
7725
7730
7735
7740
7745
7750
7755
7760
7765
7770
7775
7780
7785
7790
7795
7800
7805
7810
7815
7820
7825
7830
7835
7840
7845
7850
7855
7860
7865
7870
7875
7880
7885
7890
7895
7900
7905
7910
7915
7920
7925
7930
7935
7940
7945
7950
7955
7960
7965
7970
7975
7980
7985
7990
7995
8000
8005
8010
8015
8020
8025
8030
8035
8040
8045
8050
8055
8060
8065
8070
8075
8080
8085
8090
8095
8100
8105
8110
8115
8120
8125
8130
8135
8140
8145
8150
8155
8160
8165
8170
8175
8180
8185
8190
8195
8200
8205
8210
8215
8220
8225
8230
8235
8240
8245
8250
8255
8260
8265
8270
8275
8280
8285
8290
8295
8300
8305
8310
8315
8320
8325
8330
8335
8340
8345
8350
8355
8360
8365
8370
8375
8380
8385
8390
8395
8400
8405
8410
8415
8420
8425
8430
8435
8440
8445
8450
8455
8460
8465
8470
8475
8480
8485
8490
8495
8500
8505
8510
8515
8520
8525
8530
8535
8540
8545
8550
8555
8560
8565
8570
8575
8580
8585
8590
8595
8600
8605
8610
8615
8620
8625
8630
8635
8640
8645
8650
8655
8660
8665
8670
8675
8680
8685
8690
8695
8700
8705
8710
8715
8720
8725
8730
8735
8740
8745
8750
8755
8760
8765
8770
8775
8780
8785
8790
8795
8800
8805
8810
8815
8820
8825
8830
8835
8840
8845
8850
8855
8860
8865
8870
8875
8880
8885
8890
8895
8900
8905
8910
8915
8920
8925
8930
8935
8940
8945
8950
8955
8960
8965
8970
8975
8980
8985
8990
8995
9000
9005
9010
9015
9020
9025
9030
9035
9040
9045
9050
9055
9060
9065
9070
9075
9080
9085
9090
9095
9100
9105
9110
9115
9120
9125
9130
9135
9140
9145
9150
9155
9160
9165
9170
9175
9180
9185
9190
9195
9200
9205
9210
9215
9220
9225
9230
9235
9240
9245
9250
9255
9260
9265
9270
9275
9280
9285
9290
9295
9300
9305
9310
9315
9320
9325
9330
9335
9340
9345
9350
9355
9360
9365
9370
9375
9380
9385
9390
9395
9400
9405
9410
9415
9420
9425
9430
9435
9440
9445
9450
9455
9460
9465
9470
9475
9480
9485
9490
9495
9500
9505
9510
9515
9520
9525
9530
9535
9540
9545
9550
9555
9560
9565
9570
9575
9580
9585
9590
9595
9600
9605
9610
9615
9620
9625
9630
9635
9640
9645
9650
9655
9660
9665
9670
9675
9680
9685
9690
9695
9700
9705
9710
9715
9720
9725
9730
9735
9740
9745
9750
9755
9760
9765
9770
9775
9780
9785
9790
9795
9800
9805
9810
9815
9820
9825
9830
9835
9840
9845
9850
9855
9860
9865
9870
9875
9880
9885
9890
9895
9900
9905
9910
9915
9920
9925
9930
9935
9940
9945
9950
9955
9960
9965
9970
9975
9980
9985
9990
9995
10000
10005
10010
10015
10020
10025
10030
10035
10040
10045
10050
10055
10060
10065
10070
10075
10080
10085
10090
10095
10100
10105
10110
10115
10120
10125
10130
10135
10140<br

100



200

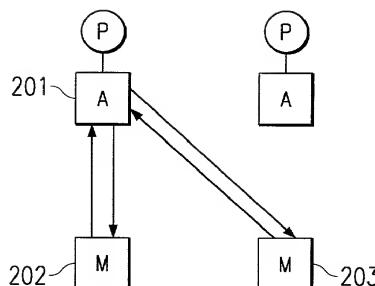
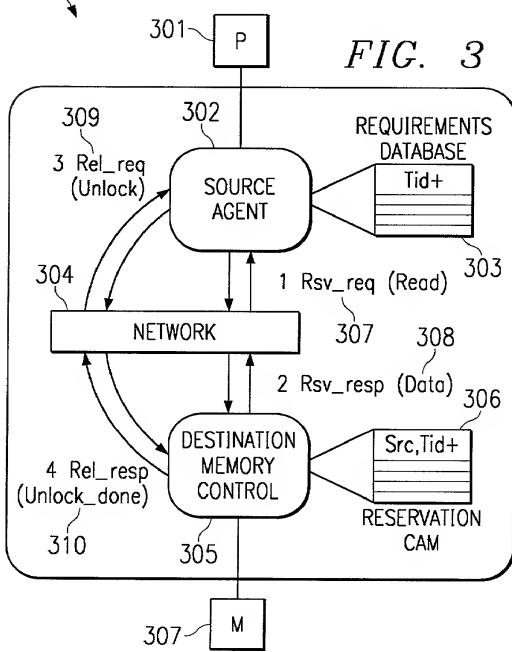


FIG. 1

FIG. 2

300



800

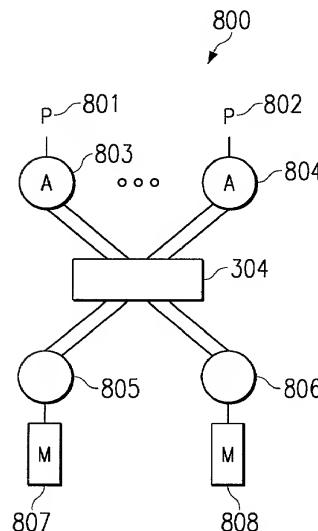


FIG. 8

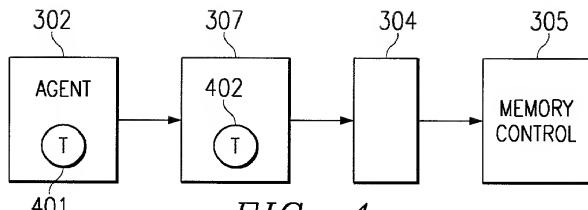


FIG. 4

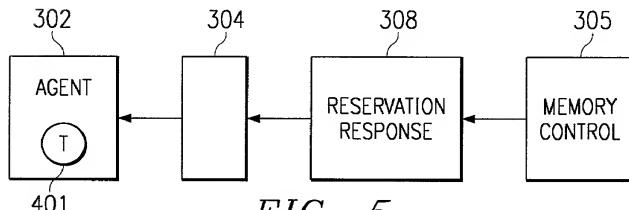


FIG. 5

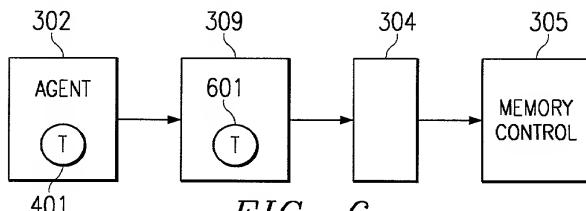


FIG. 6

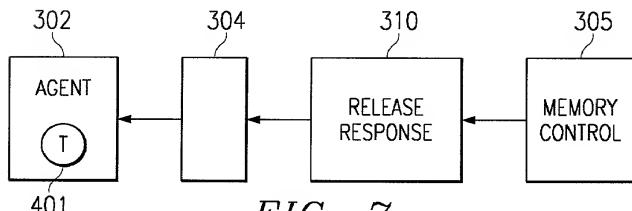


FIG. 7

DECLARATION AND POWER OF ATTORNEY
FOR PATENT APPLICATIONATTORNEY DOCKET NO. 10992150-1

As a below named inventor, I hereby declare that:

My residence/post office address and citizenship are as stated below next to my name;

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

PROTOCOL FOR INSURING EXACTLY ONCE SEMANTICS OF TRANSACTIONS ACROSS AN UNORDERED, UNRELIABLE NETWORK

the specification of which is attached hereto unless the following box is checked:

() was filed on _____ as US Application Serial No. or PCT International Application Number _____ and was amended on _____ (if applicable).

I hereby state that I have reviewed and understood the contents of the above-identified specification, including the claims, as amended by any amendment(s) referred to above. I acknowledge the duty to disclose all information which is material to patentability as defined in 37 CFR 1.56.

Foreign Application(s) and/or Claim of Foreign Priority

I hereby claim foreign priority benefits under Title 35, United States Code Section 119 of any foreign application(s) for patent or inventor(s) certificate listed below and have also identified below any foreign application for patent or inventor(s) certificate having a filing date before that of the application on which priority is claimed:

COUNTRY	APPLICATION NUMBER	DATE FILED	PRIORITY CLAIMED UNDER 35 U.S.C. 119
N/A			YES: <u> </u> NO: <u> </u> YES: <u> </u> NO: <u> </u>

Provisional Application

I hereby claim the benefit under Title 35, United States Code Section 119(e) of any United States provisional application(s) listed below:

APPLICATION SERIAL NUMBER	FILING DATE
N/A	

U. S. Priority Claim

I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code Section 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, Section 1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

APPLICATION SERIAL NUMBER	FILING DATE	STATUS (patented/pending/abandoned)
N/A		

POWER OF ATTORNEY:

As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith:

Customer Number 022879Place Customer
Number Bar Code
Label here

Send Correspondence to:
HEWLETT-PACKARD COMPANY
 Intellectual Property Administration
 P.O. Box 272400
 Fort Collins, Colorado 80528-9599

Direct Telephone Calls To:
Alexander J. Neudeck
 (970) 898-4931

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full Name of Inventor: Tony M. Brewer Citizenship: U.S.A.Residence: 5227 Mariners Drive, Plano, Texas 75093Post Office Address: 5227 Mariners Drive, Plano, Texas 75093Inventor's Signature: Tony M. Brewer Date: 2/28/00

**DECLARATION AND POWER OF ATTORNEY
FOR PATENT APPLICATION (continued)**

ATTORNEY DOCKET NO. 10992150-1

Full Name of # 2 joint inventor: Gregory S. Palmer Citizenship: U.S.A.
Residence: 3012 Mason Drive, Plano, Texas 75025
Post Office Address: 3012 Mason Drive, Plano, Texas 75025
Inventor's Signature Gregory S. Palmer Date 2/28/00

Full Name of # 3 joint inventor: _____ Citizenship: _____
Residence: _____
Post Office Address: _____
Inventor's Signature _____ Date _____

Full Name of # 4 joint inventor: _____ Citizenship: _____
Residence: _____
Post Office Address: _____
Inventor's Signature _____ Date _____

Full Name of # 5 joint inventor: _____ Citizenship: _____
Residence: _____
Post Office Address: _____
Inventor's Signature _____ Date _____

Full Name of # 6 joint inventor: _____ Citizenship: _____
Residence: _____
Post Office Address: _____
Inventor's Signature _____ Date _____

Full Name of # 7 joint inventor: _____ Citizenship: _____
Residence: _____
Post Office Address: _____
Inventor's Signature _____ Date _____

Full Name of # 8 joint inventor: _____ Citizenship: _____
Residence: _____
Post Office Address: _____
Inventor's Signature _____ Date _____